

CERTIFIED COPY OF  
PRIORITY DOCUMENT

日 本 国 特 許 庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

#  
2

JC873 U.S. PTO  
09/628306  
07/28/00

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
in this Office.

願 年 月 日  
Date of Application:

1999年 7月29日

願 番 号  
Application Number:

平成11年特許願第215348号

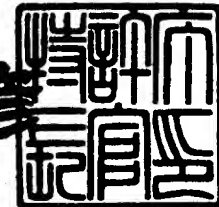
願 人  
Applicant(s):

株式会社東芝

2000年 6月29日

特許庁長官  
Commissioner,  
Patent Office

近 藤 隆 彦



出証番号 出証特2000-3051614

【書類名】 特許願

【整理番号】 A009901873

【提出日】 平成11年 7月29日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 13/00

【発明の名称】 系間情報通信システム

【請求項の数】 4

【発明者】

    【住所又は居所】 東京都府中市東芝町 1 番地 株式会社東芝府中工場内

    【氏名】 金城 守茂

【発明者】

    【住所又は居所】 東京都府中市東芝町 1 番地 株式会社東芝府中工場内

    【氏名】 廣藤 進

【発明者】

    【住所又は居所】 東京都府中市東芝町 1 番地 株式会社東芝府中工場内

    【氏名】 高桑 正幸

【特許出願人】

    【識別番号】 000003078

    【氏名又は名称】 株式会社 東芝

【代理人】

    【識別番号】 100058479

    【弁理士】

    【氏名又は名称】 鈴江 武彦

    【電話番号】 03-3502-3181

【選任した代理人】

    【識別番号】 100084618

    【弁理士】

    【氏名又は名称】 村松 貞男

【選任した代理人】

【識別番号】 100068814

【弁理士】

【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 系間情報通信システム

【特許請求の範囲】

【請求項 1】 第 1 の系と第 2 の系との間で情報通信が行われる系間情報通信システムにおいて、

前記第 1 の系と前記第 2 の系との間の転送サイズが小さい場合の情報通信に供される、転送サイズが小さい場合に高速レスポンスが可能な第 1 の通信路と、

前記第 1 の系と前記第 2 の系との間の転送サイズが大きい場合の情報通信に供される、前記第 1 の通信路に比べて転送サイズが大きい場合の転送能力が大きい第 2 の通信路とを具備し、

前記第 1 及び第 2 の系は、他系と情報通信を行う場合、対象となる情報のサイズに応じて前記第 1 または第 2 の通信路のいずれか一方が切り替え使用されるように制御する主制御手段をそれぞれ備えていることを特徴とする系間情報通信システム。

【請求項 2】 前記第 1 及び第 2 の系は、

前記主制御手段が接続された内部バスと、

自系内の前記主制御手段による前記内部バスを介しての制御のもとで前記第 1 の通信路を介して他系との間で逐次情報通信を行う第 1 のインタフェース制御手段と、

自系内の前記主制御手段からの指示に応じて当該主制御手段から独立に前記第 2 の通信路を介して他系との間で指示された情報通信を行う第 2 のインタフェース制御手段とを備えていることを特徴とする請求項 1 記載の系間情報通信システム。

【請求項 3】 前記第 1 及び第 2 の系から共通にアクセス可能な少なくとも 1 つのディスク装置と、

前記第 1 及び第 2 の系と前記ディスク装置とを接続するための第 3 の通信路とを更に具備し、

前記主制御手段は、前記第 1 の通信路及び前記第 1 のインタフェース制御手段を含む第 1 のバス上に、或いは前記第 2 の通信路及び前記第 2 のインタフェース

制御手段を含む第 2 のパス上に障害が発生した場合に、前記第 3 の通信路を前記第 1 または第 2 のパスの代替パスとして使用されるように制御することを特徴とする請求項 2 記載の系間情報通信システム。

【請求項 4】 前記第 1 及び第 2 の系は、ミラードキャッシュ方式を適用するキャッシュメモリを内蔵する 2 重化コントローラであり、

前記第 2 のインタフェース制御手段は、自系内の前記主制御手段からの指示に応じて自系内の前記キャッシュメモリに格納されたデータを前記第 2 の通信路を介して他系内の前記第 2 のインタフェース制御手段により当該他系内の前記キャッシュメモリにコピーさせることを特徴とする請求項 2 記載の系間情報通信システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、第 1 の系と第 2 の系との間で情報通信が行われる、2 重化されたコントローラを備えた 2 重化コントローラ内蔵装置に代表される系間情報通信システムに関する。

【0002】

【従来の技術】

近年のディスクアレイ装置には、それぞれ異なるホスト装置と接続可能な 2 重化されたコントローラを備えた 2 重化コントローラ内蔵装置を構成しているものがある。この種のディスクアレイ装置では、各コントローラ（各系）と対応するホスト装置との間でデータ転送が可能になっている。各コントローラには、ホスト装置との間の転送データを一時記憶するためのキャッシュメモリがそれぞれ設けられている。

【0003】

従来、上記した 2 重化コントローラを内蔵するディスクアレイ装置（系間情報通信システム）では、各コントローラ（各系）内のキャッシュメモリ上のデータの保全性を高めるために、いわゆるミラードキャッシュ方式を採用している。この方式は、ホスト装置から自系（のコントローラ）で受け取ったライトデータを

他系のコントローラのキャッシュメモリにコピー（ミラー）する方式である。

【0004】

このミラードキャッシュ方式を適用するディスクアレイ装置においては、他系でのキャッシュメモリへのコピー（ライト）が完了するのを待って、つまり自系と他系のキャッシュメモリの内容の一致化が図られるのを待って、ホスト装置へのライトの完了を通知するのが一般的である。

【0005】

また、この種のディスクアレイ装置では、コントローラ間の重要な通信での信頼性向上のために、冗長なバスを特別に備えるのが一般的である。

【0006】

【発明が解決しようとする課題】

上記したように、2重化コントローラを内蔵し、ミラードキャッシュ方式を適用する従来のディスクアレイ装置（2重化コントローラ内蔵装置、系間情報通信システム）では、他系のライトが完了するまで自系のコントローラはホスト側に完了を返せないため、この2重化コントローラ間のライトデータの通信のバスがライト時の性能のボトルネックになるという問題があった。

【0007】

また、従来のディスクアレイ装置（2重化コントローラ内蔵装置、系間情報通信システム）では、コントローラ間の重要な通信での信頼性向上のために、冗長なバスを特別に備える必要があるという問題もあった。

【0008】

本発明は上記事情を考慮してなされたものでその目的は、2重化コントローラ間の情報通信に代表される2つの系間の情報通信における転送サイズの違いにより、当該系間の情報通信に使用する通信バスを使い分けることで、当該系間の情報通信の高速化を図ることができる系間情報通信システムを提供することにある。

【0009】

本発明の他の目的は、冗長バスを特別に設けなくても、バスの冗長化を図ることができる系間情報通信システムを提供することにある。

## 【0 0 1 0】

## 【課題を解決するための手段】

本発明は、第 1 の系と第 2 の系との間で情報通信が行われる系間情報通信システムにおいて、上記第 1 の系と第 2 の系との間の転送サイズが小さい場合の情報通信に供される、転送サイズが小さい場合に高速レスポンスが可能な第 1 の通信路と、上記第 1 の系と第 2 の系との間の転送サイズが大きい場合の情報通信に供される、上記第 1 の通信路に比べて転送サイズが大きい場合の転送能力が大きい第 2 の通信路とを備えると共に、上記第 1 及び第 2 の系には、他系と情報通信を行う場合、対象となる情報のサイズに応じて上記第 1 または第 2 の通信路のいずれか一方が切り替え使用されるように制御する主制御手段をそれぞれ設けたことを特徴とする。

## 【0 0 1 1】

このような構成においては、2 つの系の間で情報通信を行う場合、その情報通信の対象となる情報のサイズ（転送サイズ）に応じて使用する通信路を使い分けることにより、つまり転送サイズが小さいときには、転送サイズが小さいときのレスポンスが高速な第 1 の通信路を用い、転送サイズが大きいときには、転送能力が大きい第 2 の通信路を用いることにより、2 つの系の間的高速情報通信が実現可能となる。ここで、転送サイズの大小は情報通信の種類によって決まることが多いため、その種類（例えば、データ転送に必要なデータ転送前後の一連の手順での系間通信、つまり制御データの通信であるか、データ転送それ自体であるか）に応じて通信路を使い分けするとよい。

## 【0 0 1 2】

また、上記第 1 及び第 2 の系に、上記主制御手段（例えばマイクロプロセッサユニット）が接続された内部バスを備えた構成では、当該第 1 及び第 2 の系に、自系内の主制御手段による内部バスを介しての制御のもとで上記第 1 の通信路を介して他系との間で逐次情報通信を行う第 1 のインタフェース制御手段と、自系内の主制御手段からの指示に応じて主制御手段から独立に上記第 2 の通信路を介して他系との間で指示された情報通信を行う第 2 のインタフェース制御手段とを設けることで、転送サイズに応じて使用する通信路を使い分けての情報通信（情

報転送)を効率的に行うことができる。

【0013】

ここで、上記第1及び第2の系が、ミラードキャッシュ方式を適用するキャッシュメモリを内蔵する2重化コントローラである構成では、上記第2のインタフェース制御手段に、自系内の主制御手段からの指示に応じて自系内のキャッシュメモリに格納されたデータを上記第2の通信路を介して他系内の第2のインタフェース制御手段により当該他系内のキャッシュメモリにコピーさせる機能を持たせるならば、内部バスでのトラヒックが上がるのを防止しながら、大量のデータコピーが実現可能となる。

【0014】

また、上記第1及び第2の系から共通にアクセス可能な少なくとも1つのディスク装置と、上記第1及び第2の系とディスク装置とを接続するための第3の通信路とを備えたシステムでは、上記第1のバスまたは第2のバス上に障害が発生した場合に、上記主制御手段の制御により、上記ディスク装置アクセス用の第3の通信路が第1または第2のバスの代替バスとして使用される構成とすることにより、冗長バスを特別に用意しなくても、通信バスの冗長化が実現可能となる。

【0015】

【発明の実施の形態】

以下、本発明の実施の形態につき、系間情報通信システムが2重化コントローラ内蔵のディスクアレイ装置である場合を例に、図面を参照して説明する。

【0016】

図1は本発明の一実施形態に係る2重化コントローラ内蔵のディスクアレイ装置のブロック構成図、図2は当該ディスクアレイ装置を備えたシステムのブロック構成図である。

【0017】

図2において、100、200は、本発明に直接関係するディスクアレイ装置を接続するホスト装置（ホストコンピュータ）であり、例えばサーバ（サーバコンピュータ）をなす。ホスト装置100、200とディスクアレイ装置300（内の後述するコントローラ310、320に設けられたホストI/F313、3



2 3) とは、S C S I (Small Computer System Interface) バス、あるいはファイバチャネル等のバス 1 1 0, 2 1 0 により接続されている。

ホスト装置 1 0 0, 2 0 0 には、例えば複数のクライアント (クライアントコンピュータ) 4 0 0 がイーサネット (Ethernet) バス、リングバス等により構成されるネットワーク 5 0 0 を介して接続されている。

【 0 0 1 8 】

ディスクアレイ装置 3 0 0 は、図 1 に示すように、当該装置 3 0 0 全体を制御する 2 重化されたコントローラ 3 1 0, 3 2 0 と、ディスクアレイを構成する複数のディスク装置、例えば 2 つのハードディスク装置 (以下、HDD と称する) 3 6 0, 3 6 1 とを備えている。コントローラ 3 1 0, 3 2 0 (内の後述する HDD - I / F 3 1 4, 3 2 4) と HDD 3 6 0, 3 6 1 とは、S C S I バス等のバス 3 5 0 により接続されている。

【 0 0 1 9 】

コントローラ 3 1 0, 3 2 0 は、当該コントローラ 3 1 0, 3 2 0 の制御中枢をなす中央制御部としての M P U (Micro Processing Unit) 3 1 1, 3 2 1 と、キャッシュメモリ部 3 1 2, 3 2 2 とを備えている。

【 0 0 2 0 】

キャッシュメモリ部 3 1 2, 3 2 2 は、ホスト装置 1 0 0, 2 0 0 から転送されたデータ (ライトデータ)、及び HDD 3 6 0, 3 6 1 から読み出されたデータ (リードデータ) を一時記憶するキャッシュメモリ 3 1 2 a, 3 2 2 を内蔵する。このキャッシュメモリ 3 1 2 a, 3 2 2 a は、ミラードキャッシュを実現するために、当該キャッシュメモリ 3 1 2 a, 3 2 2 a (を内蔵したキャッシュメモリ部 3 1 2, 3 2 2) を持つコントローラ 3 1 0, 3 2 0 と対をなすもう一方のコントローラ 3 2 0, 3 1 0 からのデータ、つまり他方 (他系) のコントローラ 3 2 0, 3 1 0 からのデータがコピー可能な領域を有する。またキャッシュメモリ部 3 1 2, 3 2 2 は、I / F (インタフェース) として、後述する内部バス 3 1 7, 3 2 7 との接続用と、コントローラ間データ転送用 (コントローラ間データ転送に用いるバス 3 4 0 との接続用) との 2 ポートを有する。

【 0 0 2 1 】

コントローラ 310, 320 はまた、ホスト装置 100, 200 と接続するバス 110, 120 の制御部をなすインタフェース制御部、つまりホスト装置 100, 200 とのインタフェース制御部（以下、ホスト I/F と称する）313, 323 と、HDD 360, 361 とのインタフェース制御部（以下、HDD-I/F と称する）314, 324 とを備えている。

#### 【0022】

コントローラ 310, 320 はまた、当該コントローラ 310, 320 間のデータ転送制御を司り、データ転送サイズが小さい場合に高速レスポンスが可能なコントローラ間データ転送制御部（以下、コント間 I/F と称する）315, 325 と、当該コントローラ 310, 320 間のデータ転送制御を司り、データ転送サイズに影響されずに一定のレスポンスが可能なコントローラ間データ転送制御部（以下、コント間 I/F と称する）316, 326 とを備えている。コント間 I/F 315, 325 は内部バス 317, 327 を介して MPU 311, 321 と接続され、コント間 I/F 316, 326 はキャッシュメモリ部 312, 322 のコントローラ間データ転送用ポートと接続されている。

#### 【0023】

コントローラ 310, 320 は更に、PCI バス (Peripheral Component Interconnect Bus) に代表される内部バス 317, 327 を備えている。この内部バス 317, 327 には、コントローラ 310, 320 内の各モジュール（ここでは、MPU 311, 321 と、キャッシュメモリ部 312, 322 と、ホスト I/F 313, 323 と、HDD-I/F 314, 324 と、コント間 I/F 315, 325 の各モジュール）が接続されている。

#### 【0024】

コントローラ 310, 320 内のコント間 I/F 315, 325 は、データ転送サイズが小さいデータの転送に適したシリアルバスに代表されるバス 330 により相互接続され、コント間 I/F 316, 326 は、データ転送サイズが大きいデータの転送に適した非同期転送型のバス 340 により相互接続されている。従来のディスクアレイ装置には、コント間 I/F 316, 326 と、当該コント間 I/F 316, 326 を接続するバス 340 とが存在しないことに注意された

い。このため従来は、コントローラ 3 1 0, 3 2 0 間のデータ転送は、MPU 3 1 1, 3 1 2 の制御により全て内部バス 3 1 7, 3 2 7 を介して行う必要がある。

#### 【0 0 2 5】

次に、図 2 のシステムに適用される図 1 の構成のディスクアレイ装置 3 0 0 の動作について、ホスト装置 1 0 0 からコントローラ 3 1 0 に対して要求されたデータのライト時に、そのコピーデータをもう一方のコントローラ 3 2 0 内のキャッシュメモリ 3 2 2 に格納する場合を例に、図 3 及び図 4 を参照して説明する。

#### 【0 0 2 6】

今、ホスト装置 1 0 0 からディスクアレイ装置 3 0 0 内のコントローラ 3 1 0 に対し、バス 1 1 0 を介してデータのライト要求が送られたものとする。コントローラ 3 1 0 が、このホスト装置 1 0 0 からのライト要求をホスト I / F 3 1 3 を介して受け付けると、ホスト装置 1 0 0 からコントローラ 3 1 0 へのライトデータの転送が開始される。

#### 【0 0 2 7】

すると、ホスト装置 1 0 0 からバス 1 1 0 を介して転送されるデータは、コントローラ 3 1 0 内のホスト I / F 3 1 3 で受け取られて、ホスト I / F 3 1 3 → 内部バス 3 1 7 → キャッシュメモリ部 3 1 2 のバスを通過して、そのコントローラ 3 1 0 のキャッシュメモリ部 3 1 2 に内蔵されたキャッシュメモリ 3 1 2 に格納される。本実施形態では、ホスト装置 1 0 0 からのライト要求で指定されたディスクアレイ（内の HDD）のデータがキャッシュメモリ 3 1 2 a にも存在する（ライト時の）キャッシュヒットの場合には、当該キャッシュメモリ 3 1 2 a 内の該当するデータが、ホスト装置 1 0 0 からのライトデータに更新される。このキャッシュメモリ 3 1 2 a に格納されたライトデータは、ホスト装置 1 0 0 からのライト要求で指定されたディスクアレイ（内の HDD）に対する遅延書き込みの対象となる。

#### 【0 0 2 8】

コントローラ 3 1 0 内の MPU 3 1 1 は、当該コントローラ 3 1 0（自系）のキャッシュメモリ 3 1 2 に格納されたのと同じのライトデータをコントローラ 3

20（他系）のキャッシュメモリ322にコピー（ミラー）するために、コントローラ320が停止していないか否かを、内部バス317を介してのコント間I/F315を用いたコントローラ間通信（制御データの授受）で確認する。即ちMPU311は、コントローラ320内のMPU321への問い合わせを、図3に示すように、MPU311→内部バス317→コント間I/F315→バス330→コント間I/F325→内部バス327→MPU321のバス31を用いて通知する。

【0029】

コントローラ320内のMPU321は、コントローラ310内のMPU311からの問い合わせを（コント間I/F325から内部バス327を介して）受け取ると、その問い合わせとは逆のパスで、つまり図3に示すように、MPU321→内部バス327→コント間I/F325→バス330→コント間I/F315→内部バス317→MPU311のバス32で、コントローラ310内のMPU311へ正常である旨の応答を返す。

【0030】

MPU311は、コントローラ320内のMPU321からの正常応答を受け取ると、先に格納したデータのキャッシュメモリ322へのコピーを、図4に示すように、MPU311→内部バス317→キャッシュメモリ部312のバス41を用いてキャッシュメモリ部312に指示する。

【0031】

するとキャッシュメモリ部312は、指示されたデータを、コント間I/F316を用いたデータ転送により、図4に示すように、キャッシュメモリ312a→コント間I/F316→バス340→コント間I/F326→キャッシュメモリ322aのバス42を用いて、コントローラ320内のキャッシュメモリ322a（に確保されているコピー領域）にコピーさせる。このデータ転送の最後では、データ転送結果を示す転送ステータスが（コントローラ320内の）キャッシュメモリ部322のコントローラ間データ転送用I/F（ポート）から上記バス42と逆のパスで、（コントローラ310内の）キャッシュメモリ部312のコントローラ間データ転送用I/F（ポート）に返される。

## 【0032】

キャッシュメモリ部 3 1 2 は、上記転送ステータスからデータ転送の完了を判定すると、内部バス 3 1 7 との接続用 I / F (ポート) により、キャッシュメモリ 3 2 2 a への転送 (コピー) が完了したことを、図 4 に示すように、キャッシュメモリ部 3 1 2 → 内部バス 3 1 7 → MPU 3 1 1 のバス 4 3 を用いて MPU 3 1 1 に通知する。

## 【0033】

MPU 3 1 1 は、このキャッシュメモリ 3 2 2 a への転送 (コピー) 完了通知を受け取って、キャッシュメモリ 3 1 2 a からキャッシュメモリ 3 2 2 a へのコピー (ミラー) の完了を確認すると、ホスト装置 1 0 0 から要求されたデータライトの完了を判定し、MPU 3 1 1 → 内部バス 3 1 7 → ホスト I / F 3 1 3 → バス 1 1 0 → ホスト装置 1 0 0 のバスによりホスト装置 1 0 0 にライト完了通知を返す。なお、キャッシュメモリ 3 1 2 a からキャッシュメモリ 3 2 2 a へのコピー (ミラー) が完了した旨の通知は、転送サイズが小さいため、前記バス 3 2 を用いてコントローラ 3 1 0 内の MPU 3 1 1 送るようにしても構わない。

## 【0034】

このように本実施形態では、コントローラ 3 1 0, 3 2 0 (内の MPU 3 1 1, 3 2 1) 間の上記問い合わせ及び当該問い合わせに対する応答、つまりコントローラ間通信 (制御データの授受) には、コント間 I / F 3 1 5, 3 2 5 及びバス 3 3 0 を含むバス 3 1 (3 2) を用い、キャッシュメモリ 3 1 2 a, 3 2 2 a 間のデータ転送 (データコピー) には、コント間 I / F 3 1 6, 3 2 6 及びバス 3 4 0 を含むバス 4 2 を用いるようにしている。

## 【0035】

ここで、コントローラ間通信は、大容量のデータの転送が行われるキャッシュメモリ 3 1 2 a, 3 2 2 a 間のデータ転送 (データコピー) に比べて、その転送データ量が著しく少ない。

## 【0036】

次に、バス 3 1 の中心をなすバス 3 3 0 は、本実施形態のようにシリアルバスで構成されている場合、データ転送開始までの手順と、データ転送終了後の手順

に（伴う制御データの授受に）要する時間は、非同期転送型のバス 3 4 0 と異な  
って少ない。つまり転送のための手順のオーバーヘッドが小さい。一方、データ転  
送時間は、データ量が多いほど長く、少ないほど短くなる。したがってバス 3 3  
0 は、上記のコントローラ間通信のように、データ量が少ないデータ転送の場合  
に高速レスポンスが期待できる。

## 【 0 0 3 7 】

ところが従来は（コント間 I / F 3 1 6, 3 2 7 及びバス 3 4 0 が存在しない  
ため）、データ転送量が少ないコントローラ間通信は勿論、データ転送量が多い  
キャッシュメモリ間データ転送（データコピー）の場合にも、（コント間 I / F  
3 1 5, 3 2 5 及びバス 3 3 0 を含む）上記バス 3 1（3 2）を用いていた。し  
かも、このバス 3 1（3 2）には各コントローラ 3 1 0, 3 2 0 の内部バス 3 1  
7, 3 2 7 が含まれている。したがって、データコピーのために内部バス 3 1 7  
, 3 2 7 を使用することでバストラヒックが上がり、加えて M P U 3 1 1, 3 2  
1 の処理が必要となることから、コピー性能、更にはホスト装置 1 0 0, 2 0 0  
からの要求処理性能が低下する虞があった。

## 【 0 0 3 8 】

これに対して本実施形態では、キャッシュメモリ部 3 1 2, 3 2 2 に接続され  
たコント間 I / F 3 1 6, 3 2 6 を設け、当該コント間 I / F 3 1 6, 3 2 6 に  
より、（内部バス 3 1 7, 3 2 7 から独立した）バス 3 4 0 を中心とするバス 4  
1 を介して、キャッシュメモリ部 3 1 2, 3 2 2 内のキャッシュメモリ 3 1 2 a  
, 3 1 2 b 間の（データコピーのための）データ転送を M P U 3 2 1, 3 2 2 か  
ら独立に行うようにしている。ここでは、データコピーのために、コントローラ  
3 1 0, 3 2 0 内の各種モジュールが接続されている内部バス 3 1 7, 3 2 7 を  
使用せず、M P U 3 2 1, 3 2 2 の処理も不要なことから、上述の性能低下を低  
減できる。

## 【 0 0 3 9 】

ここで、バス 3 4 0 が本実施形態のように同期転送型の場合、データ転送開始  
までとデータ転送終了後の手順が複雑であり、バス 3 3 0 と異なって転送のため  
の手順のオーバーヘッドが大きい。一方、バス 3 3 0 経由でのデータコピー等、一

定の転送サイズの範囲でのデータ転送に要する時間は、上記手順に要する時間に比べて短い。したがって、コント間 I / F 3 1 6, 3 2 6 によりバス 3 4 0 を経由してのデータ転送では、一定の転送サイズの範囲では、転送サイズに関係なく、データ転送のための手順に要する時間でほぼ決まる一定のレスポンスを持つ。

## 【 0 0 4 0 】

また、上記のように、データコピーに内部バス 3 1 7, 3 2 7、コント間 I / F 3 1 5, 3 2 5 及びバス 3 3 0 を含むバス 3 1 ( 3 2 ) を使用しない構成では、つまりバス 3 1 ( 3 2 ) を、コントローラ間通信のようにデータ量が少ないデータ転送のみに使用する構成では、価格的な制約のもとで、バス 3 3 0 の信号ビット幅を少なくし (ハードウェア量を減らし)、代わりに高速信号線を用いることが可能となり、コントローラ間通信等を高速に処理することができる。

## 【 0 0 4 1 】

ところで、図 2 の構成のディスクアレイ装置 3 0 0 において、上記バス 3 1 ( 3 2 ) の中心をなすコント間 I / F 3 1 5、バス 3 3 0、コント間 I / F 3 2 5、或いは上記バス ( 4 2 ) の中心をなすコント間 I / F 3 1 6、バス 3 4 0、コント間 I / F 3 2 6 のいずれかに障害が発生することがあり得る。

## 【 0 0 4 2 】

そこで本実施形態では、これらのバス上のいずれかの要素に障害が発生した場合の代替バスとして、HDD - I / F 3 1 4、バス 3 5 0 及び HDD - I / F 3 2 4 のバスが使用可能なようになっている。この HDD - I / F 3 1 4、バス 3 5 0 及び HDD - I / F 3 2 4 のバス (代替バス) を実現するには、例えばコントローラ 3 1 0 であれば、内部バス 3 1 7 上の通信に使用するデバイスの指定を、コント間 I / F 3 1 5 或いはキャッシュメモリ部 3 1 2 のコントローラ間データ転送用 I / F (ポート) から HDD - I / F 3 1 4 へ、コントローラ 3 2 0 であれば、内部バス 3 2 7 上の通信に使用するデバイスの指定をコント間 I / F 3 2 5 或いはキャッシュメモリ部 3 2 2 のコントローラ間データ転送用 I / F (ポート) から HDD - I / F 3 2 4 へ変更すればよい。

## 【 0 0 4 3 】

なお、コントローラ 3 1 0, 3 2 0 のホスト I / F 3 1 3, 3 2 3 に、ホスト

装置との接続ポートを2つ持たせ、ホスト装置100, 200を、図5に示すように、バス110, 210に相当するバス120, 220を介してコントローラ320, 310のホストI/F323, 313に接続するとよい。

#### 【0044】

このような構成では、例えばコントローラ310自身の障害またはホスト装置100と当該コントローラ310との間の通信バスの障害が発生した場合、ホスト装置100はバス120を用いることで、他系のコントローラ320を切り替え使用することができる。同様にコントローラ320自身の障害またはホスト装置200と当該コントローラ320との間の通信バスの障害が発生した場合には、ホスト装置200はバス220を用いることで、他系のコントローラ310を切り替え使用することができる。

#### 【0045】

また図5の構成において、ホスト装置100, 200は、相手に障害が発生した場合に相手の代替処理を行うようになっている。そのためホスト装置100, 200は互いに障害発生を監視している。この監視は、ネットワーク500を介して行うことも可能であるが、トラフィックと信頼性を考慮するならば、図5に示すようにホスト装置100, 200間を接続するイーサネットバス等のバス600を設け、当該バス600を介して監視するとよい。なお、ホスト装置100, 200は、互いに処理を行う構成でも、一方が待機系となる構成のいずれでも構わない。

#### 【0046】

以上は、本発明を、2重化されたコントローラ間の情報通信の高速化のために、その情報通信の対象となる情報のサイズ（の違いを反映する情報通信の種類、つまりデータ転送前後の一連の手順でのコントローラ間通信であるか、データ転送それ自体であるか）に応じて2種の通信バスを切り替え使用する2重化コントローラ内蔵のディスクアレイ装置に実施した場合について説明したが、これに限るものではない。本発明は、相互に情報通信を行う2つの系の装置を備えた系間情報通信システムにおける、系間の情報通信の高速化にも適用可能である。

#### 【0047】



## 【発明の効果】

以上詳述したように本発明によれば、2重化されたコントローラ間の情報通信に代表される2つの系の間の情報通信における転送サイズの違い（を反映する情報通信の種類の違い）により、当該系間の情報通信に使用する通信バスを使い分けるようにしたので、当該系間の情報通信の高速化を図ることができる。特に、ミラードキャッシュ方式を適用するキャッシュメモリを内蔵する2重化コントローラ間の情報通信を行うシステムでは、コントローラ間的高速データコピーが実現できる。

## 【0048】

また本発明によれば、第1及び第2の系から共通にアクセス可能な少なくとも1つのディスク装置と、これらを相互接続するための通信路とを備えたシステムでは、系間の情報通信に切り替え使用する通信バスの障害時には、この通信路を代替バスとして使用することにより、冗長バスを特別に設けなくても、バスの冗長化を図ることができる。

## 【図面の簡単な説明】

## 【図1】

本発明の一実施形態に係る2重化コントローラ内蔵のディスクアレイ装置のブロック構成図。

## 【図2】

図1の構成のディスクアレイ装置を備えたシステムのブロック構成図。

## 【図3】

同実施形態におけるコントローラ間通信で使用される通信バスを説明するための図。

## 【図4】

同実施形態におけるデータコピーで使用される通信バスを説明するための図。

## 【図5】

図2のシステムの変形例を示すブロック図。

## 【符号の説明】

31, 32…バス（第1のバス）

4 2 …バス (第 2 のバス)

1 0 0, 2 0 0 …ホスト装置

3 0 0 …ディスクアレイ装置 (系間情報通信システム、2 重化コントローラ内蔵装置)

3 1 0 …コントローラ (第 1 の系)

3 1 1, 3 2 1 …MPU (主制御手段)

3 1 2, 3 2 2 …キャッシュメモリ部

3 1 2 a, 3 2 2 a …キャッシュメモリ

3 1 3, 3 2 3 …ホスト I / F

3 1 4, 3 2 4 …HDD - I / F

3 1 5, 3 2 5 …コント間 I / F (第 1 のインタフェース制御手段)

3 1 6, 3 2 6 …コント間 I / F (第 2 のインタフェース制御手段)

3 1 7, 3 2 7 …内部バス

3 2 0 …コントローラ (第 2 の系)

3 3 0 …バス (第 1 の通信路)

3 4 0 …バス (第 2 の通信路)

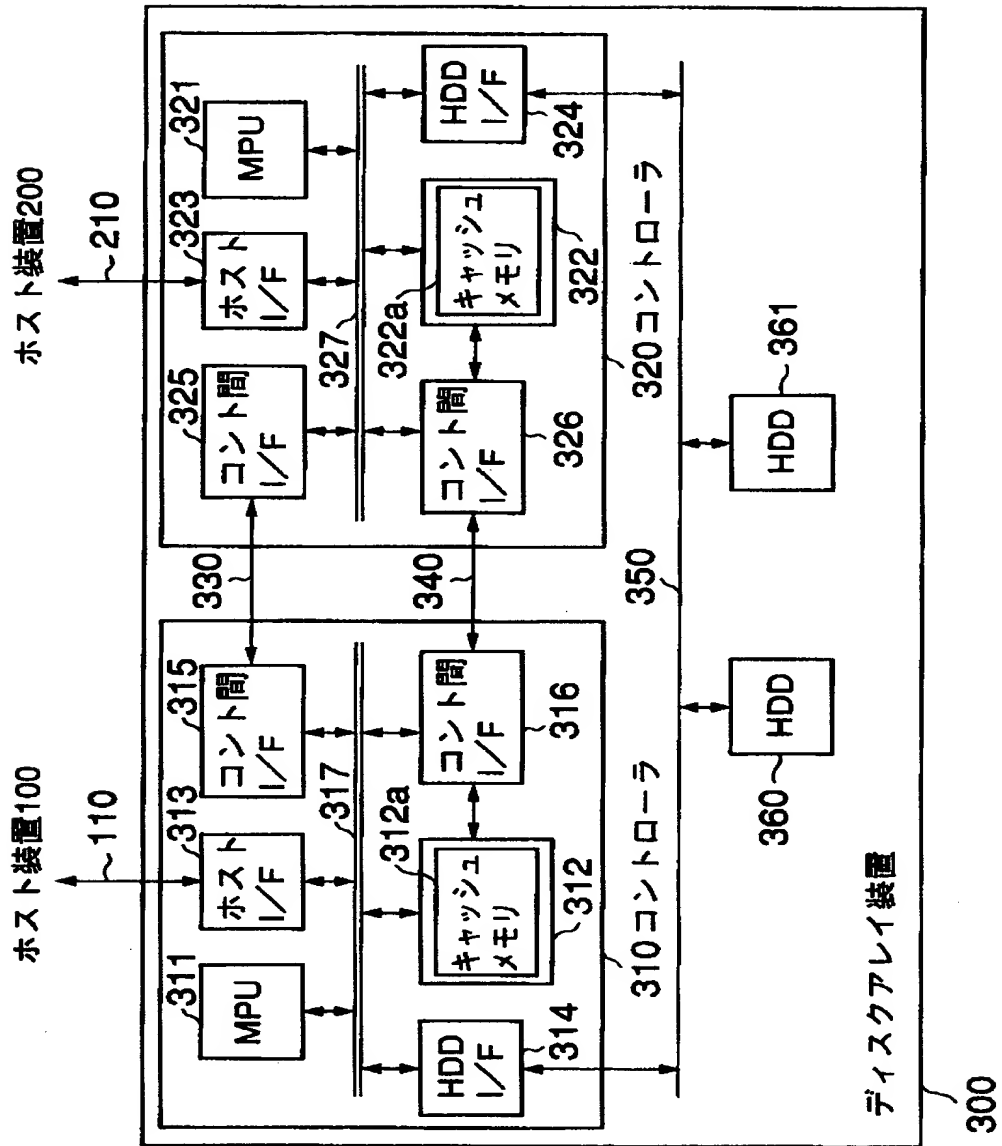
3 5 0 …バス (第 3 の通信路)

3 6 0, 3 6 1 …HDD (ディスク装置)

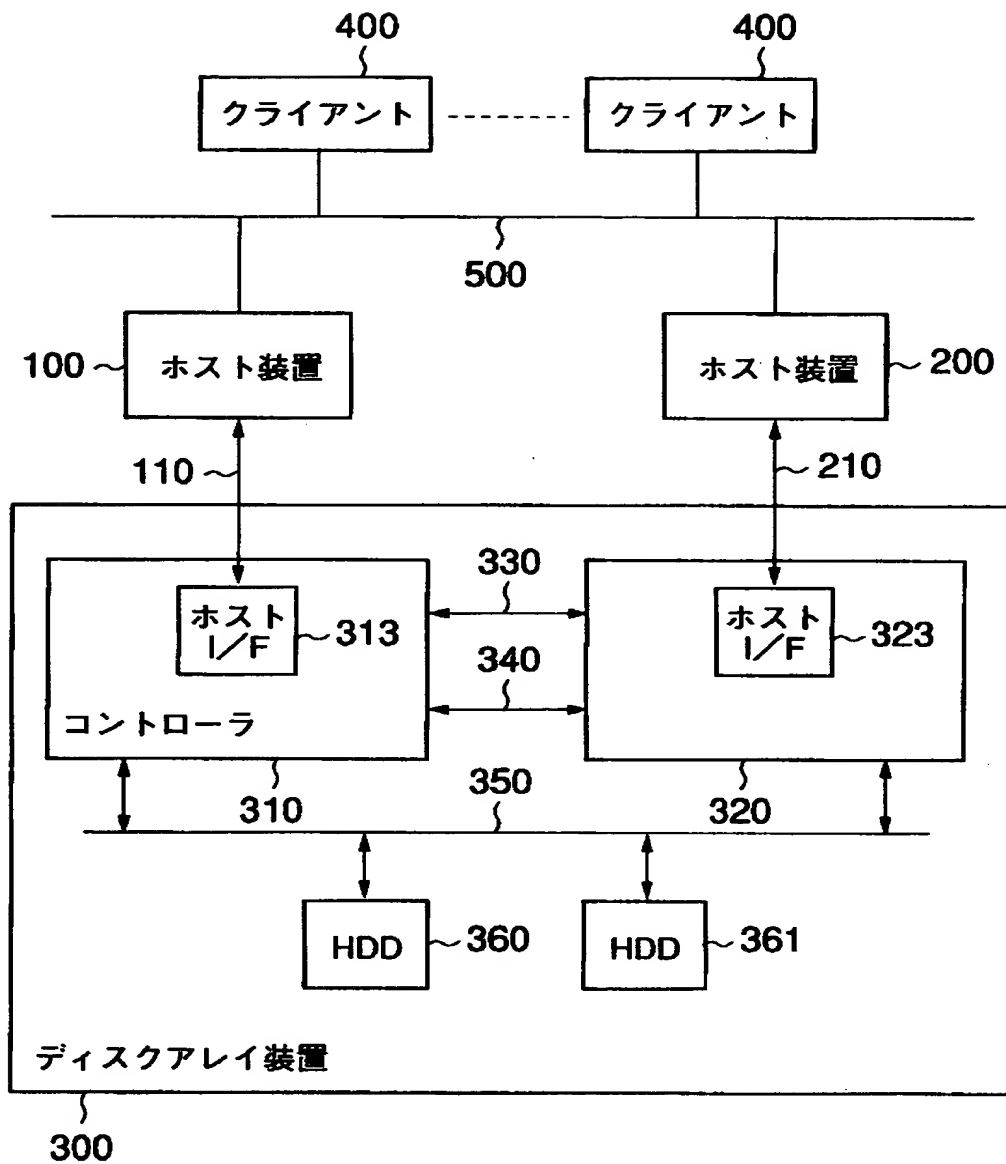
【書類名】

図面

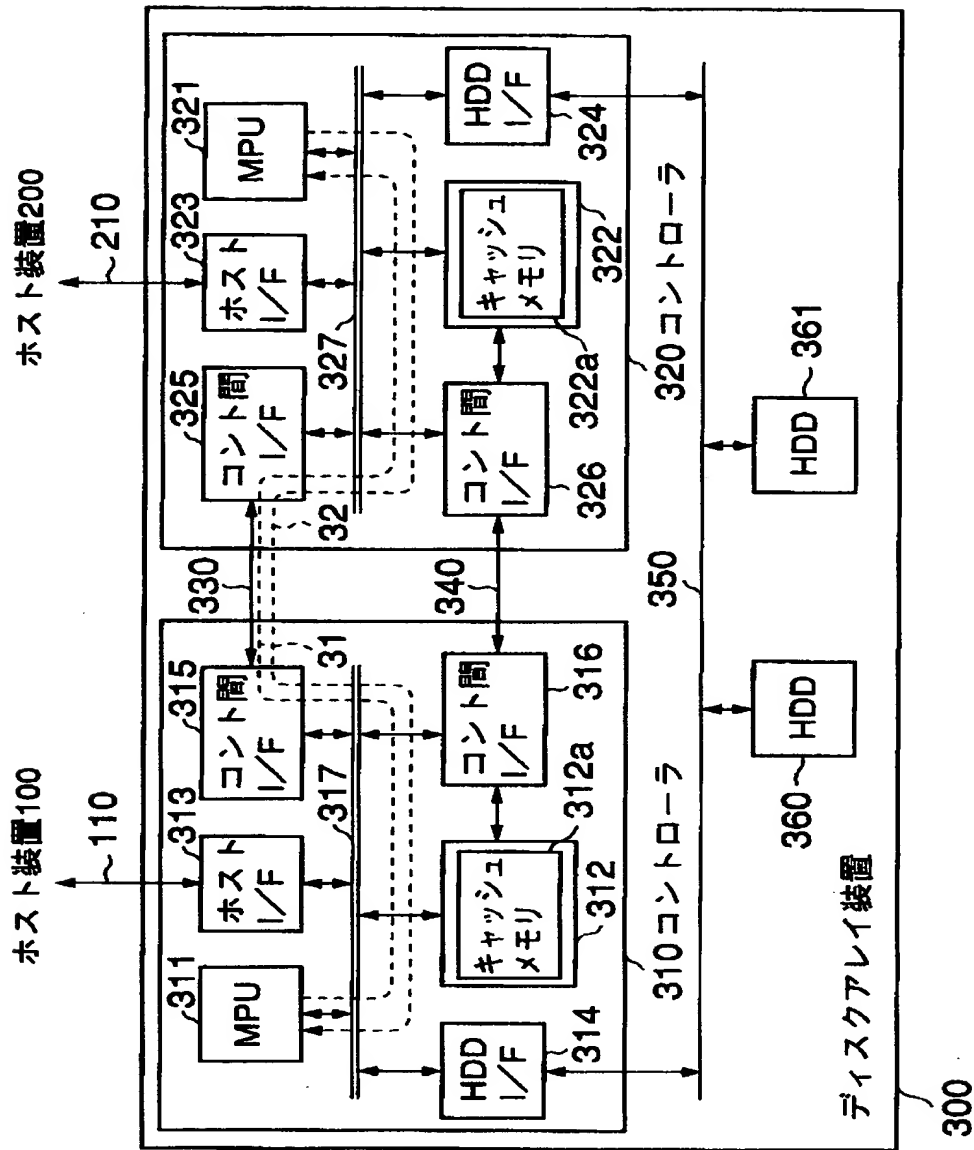
【図 1】



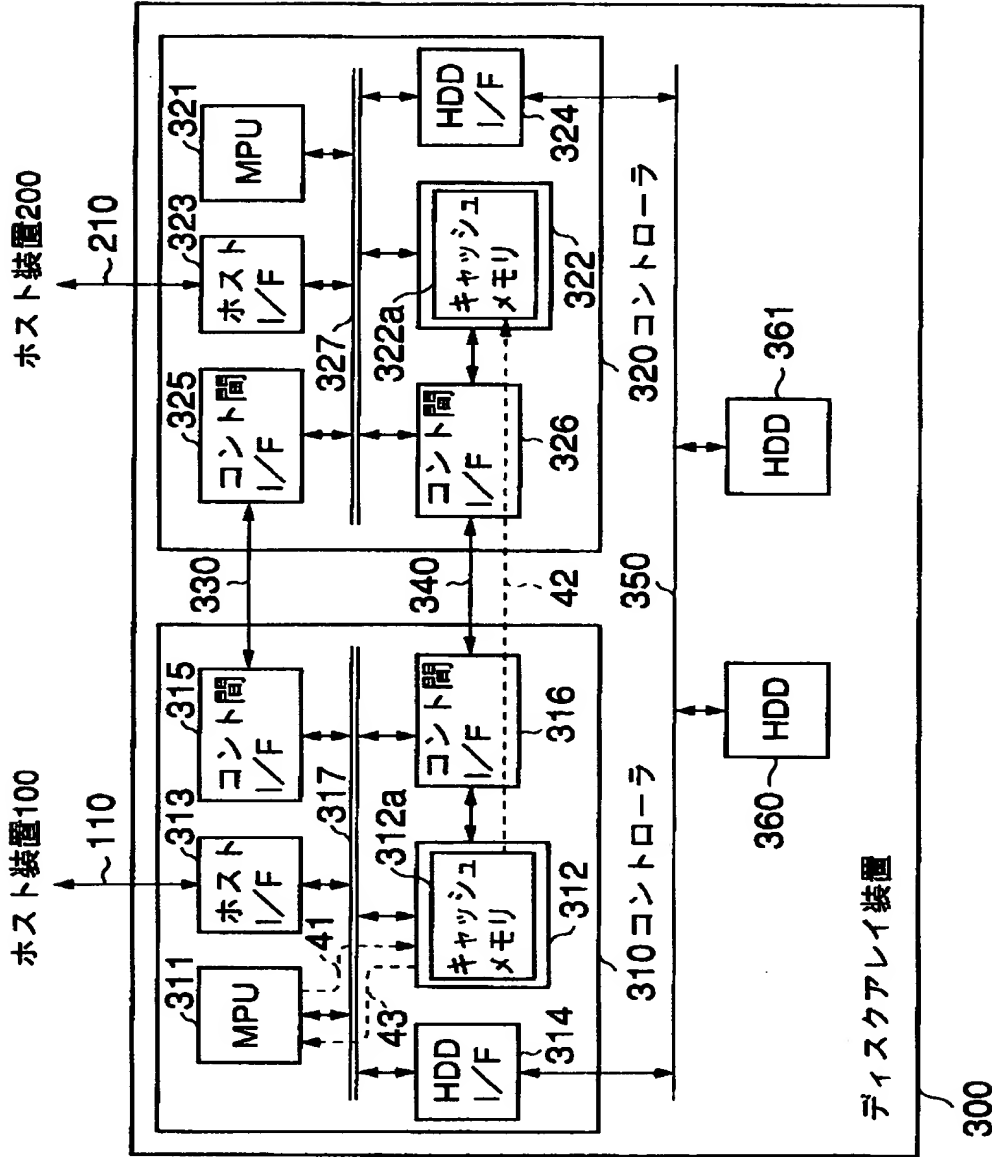
【図 2】



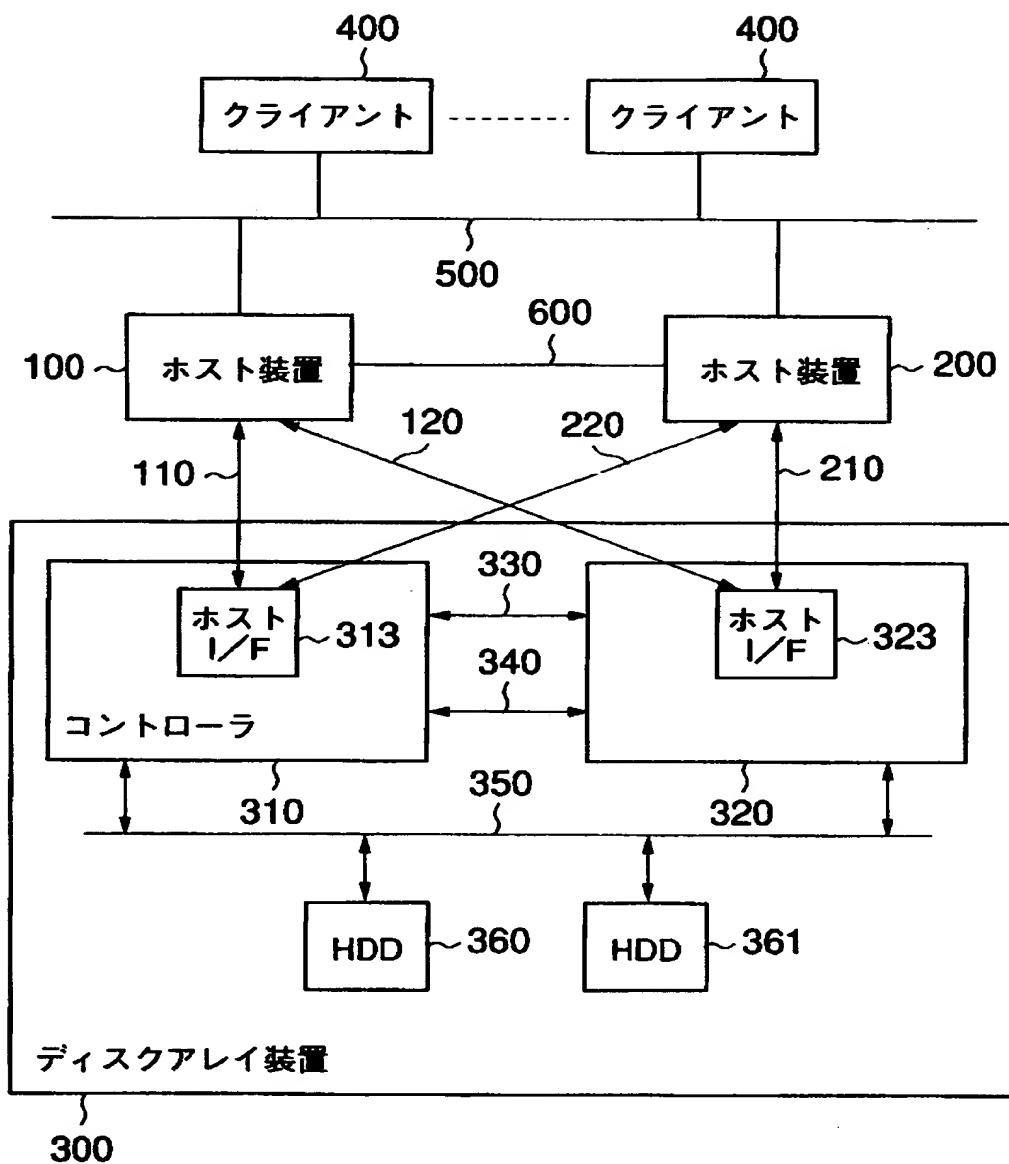
【図 3】



【図 4】



【図 5】



【書類名】 要約書

【要約】

【課題】 2重化コントローラ間の情報通信に代表される2つの系の間の情報通信における転送サイズの違いにより、当該系間の情報通信に使用する通信パスを使い分けることで、当該系間の情報通信の高速化を図る。

【解決手段】 キャッシュメモリ 3 1 2 a, 3 2 2 a を内蔵する2重化されたコントローラ 3 1 0, 3 2 0 間の通信と、キャッシュメモリ 3 1 2 a, 3 1 2 a 間のデータコピーでの転送サイズの違い（を反映する情報通信の種類の違い）により、前者の場合には、MPU 3 1 1、内部バス 3 1 7、コント間 I / F 3 1 5、バス 3 3 0、コント間 I / F 3 2 5、内部バス 3 2 7、MPU 3 2 1 のパスを、後者の場合には、キャッシュメモリ 3 1 2 a、コント間 I / F 3 1 6、バス 3 4 0、コント間 I / F 3 2 6、キャッシュメモリ 3 2 2 a のパスを、それぞれ切り替えて使用する。

【選択図】 図 1



出 願 人 履 歴 情 報

識別番号 [000003078]

1. 変更年月日	1990年 8月22日
[変更理由]	新規登録
住 所	神奈川県川崎市幸区堀川町72番地
氏 名	株式会社東芝